


Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

August 13, 2021

To: Ron S. Jarmin 
Acting Director
U.S. Census Bureau

From: Allison Plyer
Census Scientific Advisory Committee Chair

Subject: Recommendations and Comments to the Census Bureau from the
Census Scientific Advisory Committee Differential Privacy Virtual Meeting

May 25, 2021

The Census Scientific Advisory Committee (CSAC) is impressed with the Bureau's development of Differential Privacy (DP) as the core of its enhanced Disclosure Avoidance System. It represents an enormous achievement. Indeed, MIT's Technology Review recognized the Bureau's development of DP as one of 10 top Breakthrough Technologies of 2020 (along with anti-aging drugs, unhackable internet, and 7 others). MIT notes that other countries such as the U.K. and Canada are watching the Census Bureau's development of DP as are other federal agencies that may adopt this methodology following the Census Bureau's lead.¹ For these reasons, and more, it is critical that the pros and cons of this innovation are well understood and documented. CSAC also appreciates the Bureau's efforts in creating the 2010 Demonstration Products, the Detailed Summary Metrics and other updates, and the privacy-protected microdata files for evaluation by the community of users. CSAC applauds the Bureau for adapting its algorithms in response to feedback from that community.

I) **General Communications and Engagement about Census Products and DP**

CSAC commends the Bureau for its communications with users through blogs, webinars, newsletters and presentations about the forthcoming 2020 Census products and the evolving plans for applying DP. Serious concerns remain about how differential privacy will affect the data's fitness of use for different use cases, so communication and transparency will be key to maintaining users' trust. The response to Recommendation numbered 8 on page 13 of responses to our Fall 2020 CSAC recommendations that explained the working priority order of use cases is a good example of the types of explanations that will help users understand the Bureau's decisions as they make tradeoffs in allocating the privacy-loss budget.

¹ <https://www.technologyreview.com/10-breakthrough-technologies/2020/#differential-privacy>

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

- 1) **CSAC recommends that the Bureau continue its efforts to regularly communicate updates and engage various user groups during the decision-making process for setting the privacy-loss budget (PLB) and its allocation. CSAC also recommends that the Bureau actively engage with the stakeholder community** - researchers, local and state government staff, other federal agency staff, and others - as the Bureau makes final decisions about the TopDown Algorithm to release the P.L. 94-171 Redistricting Data, and going forward with remaining products.

Census Bureau Response: The Census Bureau accepts this recommendation. The Census Bureau has already communicated with the stakeholder community regarding the production settings for the 2020 Census P.L. 94-171 Redistricting Data Summary File and is planning a stakeholder engagement process for tuning of the Disclosure Avoidance System for the remaining 2020 Census data products. We will announce the schedule and milestones for that stakeholder engagement as soon as possible.

- 2) **CSAC also recommends that the Bureau communicate the factors used by the Data Stewardship Executive Policy (DSEP) Committee to set the PLB (“level of epsilon”).**

Census Bureau Response: The Census Bureau accepts this recommendation. We have already communicated the major factors considered by the DSEP in setting the PLB for the 2020 Census P.L. 94-171 Redistricting Data Summary File in our [June 9, 2021, news release](#) and will continue to communicate additional information about that and future decision-making as appropriate.

CSAC endorses the decision of the Bureau to publish plain-language explanations for users about practical implications of DP, starting with the P.L. 94-171 file and updated as new products are released.

- 3) **CSAC recommends that the Bureau publish 2020 Census data handbooks for data users targeted to different audiences (AIAN, federal agencies, data for rural areas, media, local government officials, etc.) that parallel the handbooks created for the American Community Survey.²**

Census Bureau Response: The Census Bureau accepts this recommendation. We are working with the Population Reference Bureau (PRB) on the development of a 2020 Census handbook series explaining differential privacy implementation and its impacts. The first handbook titled “An Introduction to Differential Privacy” is scheduled to be published by September 30, 2021.

² <https://www.census.gov/programs-surveys/acs/guidance/handbooks.html>

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

II) Considerations for Accuracy and PLB Allocation for Different Geographies and Special Populations

CSAC commends the Census Bureau's efforts to improve the DP-adjusted estimates for off-spine geography and to increase the PLB allocation to block-level geography for the P.L. 94-171. CSAC recognizes the importance of block-level data in constructing higher-level undefined geographies such as Congressional districts rather than their direct use. However, users continue to focus on accuracy at the block-level.

- 1) CSAC recommends that the Census Bureau publish evaluations and examples to show how error declines with the aggregation of blocks into previously undefined geographies and to demonstrate that biases introduced by post-processing do not accumulate into larger errors for these undefined geographies, especially in low population density regions.**

Census Bureau Response: The Census Bureau accepts this recommendation. We have already published Detailed Summary Metrics for the 2010 Census data processed through the DAS with the production settings approved by DSEP. We will also release a final set of 2010 Census Privacy-Protected Microdata Files in August 2021 to enable stakeholders to perform additional analysis of the changes we implemented to mitigate identified and potential biases.

- 2) CSAC recommends that the Bureau produce these analyses as quickly as possible, as they will contribute critical evidence for making decisions about the needed PLB for the redistricting file.**

Census Bureau Response: The Census Bureau accepts this recommendation. We have already published Detailed Summary Metrics for the 2010 Census data processed through the DAS with the production settings approved by DSEP. We will also release a final set of 2010 Census Privacy-Protected Microdata Files in August 2021 to enable stakeholders to perform additional analysis of the changes we implemented to mitigate identified and potential biases. Finally, we will release an updated version of the paper by Tommy Wright and Kyle Irimata ([Empirical Study of Two Aspects of the TopDown Algorithm Output for census.gov](#)) that uses the August 2021 PPMF to reanalyze the redistricting use case.

CSAC appreciates that the Bureau accepted our Fall 2020 recommendation that they should make "readily available tools for extrapolating from 2010 demonstration metrics to 2020 use cases" and understands that the Bureau is researching how to produce and provide these tools.

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

- 3) CSAC recommends that the Bureau should develop a set of tools to help users understand how error properties of custom geographies vary with distance from the geographic spine.**

CSAC encourages the Bureau to practice user-centered design with the target audience in mind in developing these tools, and

Census Bureau Response: The final set of 2010 Census Privacy-Protected Microdata Files that the Census Bureau plans to release in August 2021 will be one mechanism for stakeholders to use to evaluate the impact of the DAS on custom-generated geographies. We are still researching options for providing data users with additional guidance on understanding these impacts and will brief CSAC on this research at a future date.

- 4) CSAC requests an update on the progress of the tool research and development, either to the DP working group or the full CSAC as appropriate.**

In reviewing the metrics from the latest (April 28) demonstration data, CSAC notes that the reallocation of the PLB toward optimized block groups and blocks has improved the estimates (reduced error) for small off-spine geographies (such as incorporated places and minor civil divisions below 5,000 in population) and for census blocks. At the same time, however, the measures of error at the tract level have increased substantially (4-fold when November 2020 is compared to April 28 with an overall PLB of 12.2), to the point where tracts seem out of line with other on-spine geographies.

Census Bureau Response: The Census Bureau accepts this recommendation. We are still researching options for providing data users with additional guidance on understanding the impact of privacy protections on different geographies and will brief CSAC (either the DP working group or the full committee, as appropriate) on this research at a future date.

- 5) CSAC recommends that the Census Bureau consider whether this reduction in the quality of tract-level estimates represents an acceptable trade-off for the improvements achieved elsewhere or whether some reallocation of PLB to improve the quality of estimates for census tracts is merited.**

In addition, errors associated with school districts have increased and CSAC has not seen any metrics associated with accuracy of data on children under 18 years old. Because the data on population under 18 will be set by the data in the PL file, understanding implications of final decisions for the PL file will impact data on children to-be released in the DHC file.

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

Census Bureau Response: The Census Bureau has accepted this recommendation. When determining the production settings for the 2020 Census Redistricting Data Summary File, DSEP allocated substantial additional privacy-loss budget to statistics at the block-group level and above to address these concerns. Adding privacy-loss budget to the voting age variable improved the accuracy for both categories (0 to 17 and 18+). The August 2021 PPMF will allow users to supplement the metrics in our detailed summary if they wish.

- 6) CSAC recommends that the Census Bureau produce an analysis of the errors associated with data for children under 18, particularly but not exclusively applied to school districts.**

Census Bureau Response: The Census Bureau will consider including additional analysis on errors associated with data for children to the Detailed Summary Metrics in the context of stakeholder engagement to inform tuning of the TopDown Algorithm for production of the Demographic and Housing Characteristics File scheduled for a later release.

- 7) CSAC recommends that the Bureau consider strategies for ensuring especially high accuracy of blocks with prisons, especially where incarcerated people make up the entire population.**

Census Bureau Response: The Census Bureau has accepted this recommendation. The TopDown Algorithm's optimized geographic post-processing hierarchy specifically isolates group quarters facilities by type into their own optimized block groups to improve accuracy of data for the group quarters population by GQ type.

CSAC recognizes that numbers of inmates are publicly published by other state and federal agencies. This raises the challenge that inconsistencies between decennial census counts and numbers of inmates published elsewhere could erode data users' trust. It also raises the opportunity for the Bureau to draw on external numbers to improve decennial accuracy in these blocks. Moreover, blocks with prison populations are especially important to identify clearly and count accurately for redistricting uses.

- 8) CSAC recommends the Bureau investigate ways to keep prison block information as accurate as possible or at least consistent with data released by other sources, such as the Department of Justice. CSAC also recommends the Bureau explore using externally published numbers on the number of people in prisons in the post-processing process in order to maintain consistency across data published from various sources, to increase data user trust, and to maintain accuracy without using the privacy loss budget or compromising privacy protection.**

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

Census Bureau Response: While the Census Bureau thanks CSAC for this recommendation, the Census Bureau must reject the recommendation. The goal of a decennial census is to enumerate the population and thus the Census Bureau does not control its population totals to external sources.

- 9) CSAC recommends that the Census Bureau be mindful of and report on other special populations (such as those living in group quarters beyond the prison population, like college dorms) for which there may be other data sources that conflict with 2020 Census results and for which post-processing adjustments might be useful.

Census Bureau Response: The Census Bureau has accepted this recommendation. The changes to the TopDown Algorithm's geographic post-processing hierarchy, which isolates group quarters facilities by type into their own optimized block groups (also by GQ type), was incorporated into the algorithm to address this concern.

III) Summary and Use Case Metrics

Quality metrics at a finer scale are needed to help stakeholders, particularly in less populated places, understand the impact of DP on fitness of use for cases such as distributions of government funding or planning for community services. The previous Census Bureau response to this recommendation - that users can create their own new metrics from privacy-protected microdata files - is not reasonable for the vast majority of users.

- 1) CSAC recommends that the Bureau include more information about the range and pattern of error in releases of the Detailed Summary Metrics for future sprint cycles.

For example, the current MALPE (Mean Algebraic Percent Error) statistic could be split into the average negative relative error and the average positive relative error, rather than combining the two. Other statistics might report the number of blocks in more detailed categories than the current ones of percent error greater than 5 percent (such as 5 to 10 percent, 10 to 20 percent, greater than 20 percent). The ranges could also be split at scientifically meaningful thresholds or at variable-specific cut points, like the quintiles of the distribution. Another option is to include the range from lowest to highest percent error. CSAC understands that this is additional work for limited Census Bureau staff time but believes these additions are needed for users to be able to evaluate how to appropriately use the privacy-protected data.

CSAC appreciates the Census Bureau's continued inclusion of "impossible and improbable results" among its use case metrics. The most recent (April 28) metrics show that four of the eight impossible/improbable results have been reduced substantially (one to zero) by the combination of changes to the TopDown algorithm and an increase in the PLB. However, four of the results have changed little in frequency since the November 2020 release. Of particular note, 10.76 percent of census blocks with at least one occupied

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

housing unit have a total household population of zero, and 21.90 percent of all blocks have 100 percent occupancy in the DP estimates but not the published census data. CSAC is concerned that these cases are especially problematic for users and detract from the Bureau's messaging about the impact of DP on the quality and usability of the data.

Census Bureau Response: The Census Bureau accepts this recommendation. We are identifying additional metrics for assessing outliers, bias, and patterns of error for inclusion in future releases of the Detailed Summary Metrics.

- 2) CSAC recommends that the Census Bureau explicitly address these anomalies in its application of DP and its communication regarding such findings.**

Census Bureau Response: The Census Bureau accepts this recommendation. We are identifying additional metrics for assessing outliers, bias, and patterns of error for inclusion in future releases of the Detailed Summary Metrics. We will communicate those additions to our stakeholders as they are added to the Detailed Summary Metrics.

IV) Helping Users Understand Implications of Post-Processing

CSAC appreciates the Census Bureau's attention to improving its post-processing adjustments in order to reduce the error that these adjustments introduce into census data. CSAC is concerned that post-processing error is still large and that it may help to explain why the substantial increase in PLB from 4.5 to 12.2 did not produce larger reductions in the various error metrics (most were reduced by around one-half).

- 1) CSAC recommends that the Census Bureau communicate to users whatever success has been achieved in reducing post-processing error and provide evidence that post-processing error is not limiting in a substantial way the overall reduction in error achieved by the increase in the PLB.**

Census Bureau Response: The Census Bureau accepts this recommendation. In making its decisions about the production settings for the 2020 Census P.L 94-171 Redistricting Data Summary File, DSEP considered a range of privacy-loss budget allocations and extensive analysis of the resulting accuracy of the data. The Census Bureau is confident that the overall accuracy of the resulting data is under the direct control of the privacy-loss budget. We will consider ways to communicate these results to our stakeholders.

The noisy counts prior to post-processing in the TopDown Algorithm (TDA) give unbiased estimates, with analyzable error distributions. Having this data available would facilitate assessment of bias properties for the privacy-protected data, including potential positive biases created during post-processing, particularly in small domains where rounding up occurs to avoid negative values. A concern is that these small positive biases can accumulate as small domains are combined to create custom geographies.

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

- 2) **CSAC reasserts its earlier recommendation that the Bureau release the non-post-processed data used in TDA, which are unbiased estimates. To address the Bureau's concerns that the release of such estimates would require extensive user guidance, CSAC recommends that the Bureau consider releasing such data as a research product.**

Census Bureau Response: The Census Bureau appreciates this recommendation, but at this time, we are still evaluating whether and how we could release the non-post-processed noisy measurements file to our data users in a manner that would enable its use. Those decisions can only be made after a comprehensive privacy-risk analysis of the noisy measurements is performed after production of the 2020 Census Demographic and Housing Characteristics File.

V) Understanding and Managing Risks of Reconstruction

CSAC appreciates the Census Bureau's recognition of the privacy risks posed by exact invariants. Indeed, exact invariants impose hard constraints on the data, making reconstruction attacks both computationally easier and much more feasible. Invariants, thus, undermine the privacy technology.

- 1) **CSAC recommends that the Bureau continue to minimize the number of invariants they utilize.**

Census Bureau Response: The Census Bureau has accepted this recommendation. DSEP did not approve any additional invariants for the production of the P.L. 94-171 Redistricting Data Summary File.

The Bureau may find it useful to measure the effectiveness of the reconstruction attacks in re-identifying individuals, especially those that are different from those around them, to enable better understanding of the privacy risks at a certain PLB.

- 2) **CSAC recommends that the Bureau publish more details about the reconstruction attack, including the distribution on demographics and within-block minority status of the confirmed re-identifications.**

Census Bureau Response: The Census Bureau accepts this recommendation. We are preparing a full technical report for public release.

CSAC also recognizes that the reconstruction attack performed by the Bureau is one approach to re-identification, which uses a subset of the tables released by the Bureau. More sophisticated attacks will likely be able to go farther than the reconstruction attack reported by the Bureau. A more sophisticated attacker could reconstruct not just the CEF but the household structures. As an example, a list of same-sex households with children published in 2022 as based on the Decennial data from the DHC file, even if it is half wrong, would very likely discourage many same-sex households from responding to the future

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

Census Bureau surveys (e.g. the American Community Survey in the 2020s and the decennial in 2030). Given that foreign actors are credibly accused of trying to sow doubt about our democratic institutions, it is reasonable to be concerned about those same actors doing such an attack and then publicizing it on Facebook.

An excessively large PLB may well enable attacks such as these. While it may seem natural to use the PLB as a knob and use reconstruction attacks' success rate as the measure of privacy.

- 3) CSAC recommends that PLB be selected that clearly protects privacy against a range of realistic attacks while also balancing the critical need to publish accurate data.**

Census Bureau Response: The Census Bureau has accepted this recommendation. When making their decisions about the production settings for the 2020 Census P.L. 94-171 Redistricting Data Summary File, DSEP considered the overall balance of accuracy vs. privacy protections and selected an overall level and allocation of PLB that will ensure fitness-for-use of the resulting data while also providing effective protection against known types of privacy attack, such as reconstruction-abetted reidentification attacks.

CSAC would appreciate additional evidence to determine that the PLB of 12.2 provides sufficient privacy protection.

- 4) CSAC recommends that the Bureau continue to improve their algorithms such that the PLB can be set to achieve necessary levels of accuracy, with needed levels of protection.**

Census Bureau Response: The Census Bureau has accepted this recommendation. When making their decisions about the production settings for the 2020 Census P.L. 94-171 Redistricting Data File, DSEP considered the overall balance of accuracy vs. privacy protections, and selected an overall level and allocation of PLB that will ensure fitness-for-use of the resulting data while also providing effective protection against known types of privacy attack, such as reconstruction-abetted reidentification attacks.

VI) Understanding Implications of DP on Different Use Cases

- 1) CSAC recommends that the Census Bureau continue to collect and study (internally or with collaborators) use cases across a wide range of uses, variables, geographies, and among a wide range of stakeholders both for the PL data and for subsequent data releases.**

Until substantially more analyses have been conducted, the risks of releasing 2020 Census data products to which DP has been applied (at various epsilon levels) are not known. While the Census Bureau accepted prior recommendations of the Committee to publish

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

further details on impacted geographic levels and variables (and their combination) and committed to conduct an assessment of the accuracy and trade-offs in future versions on an ongoing basis, CSAC recognizes the need for additional evaluation from a variety of data users and reiterates the importance of keeping these activities ongoing.

***Census Bureau Response:** The Census Bureau accepts this recommendation. We will be actively engaging with the stakeholder community as part of the tuning of the TopDown Algorithm for production of the 2020 Census Demographic and Housing Characteristics File scheduled for a later release, and will continue to collect, assess, and communicate about census data use cases as part of our efforts to ensure fitness-for-use of the resulting data.*

CSAC finds the work the Census Bureau has done to evaluate the redistricting use case, assessing the errors due to the TopDown Algorithm in the congressional districts created during post-2010 redistricting (Wright and Iramata, 2020), to be highly useful for demonstrating how fit for use DP data can be in the all-important redistricting use case.

- 2) CSAC recommends that the Census Bureau extend this analysis to investigate how DP would have impacted congressional districts and examples of smaller political districts such as precincts that were drawn post 2010, using the latest demonstration data.**

***Census Bureau Response:** The Census Bureau accepts this recommendation and will consider including additional analysis based on the production settings for the redistricting data. It is important to note that Census Bureau does not collect post-census precincts, so we cannot evaluate post-2010 census precincts. The Census Bureau only collects precinct level data just before each decennial cycle for use in publishing that upcoming decennial's data and then they are removed from the census geographic universe. However, we have several small jurisdiction use cases against which we could perform these types of evaluations.*

- 3) CSAC also recommends that the Census Bureau investigate how DP impacts the population under age 18 in school districts, using the latest demonstration data.**

***Census Bureau Response:** The Census Bureau will consider including additional analysis on errors associated with data for children to the Detailed Summary Metrics in the context of stakeholder engagement to inform the tuning the TopDown Algorithm for production of the Demographic and Housing Characteristics File scheduled for a later release.*

CSAC applauds the Census Bureau's release of the 4/28/2021 demonstration data, which is allowing for additional critical evaluation from data users for a variety of uses. The data released in the P.L. 94-171 file are used for many other applications beyond redistricting, including federal, state, and local government mandates and planning, that the Census

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

Bureau has not, as far as CSAC is aware, evaluated similarly to the redistricting use. Additional rigorous analysis is needed on the impacts on funding formulas for federal agencies and Congressional staffers and on the impacts on legal mandates and regulatory practices, including protections for civil rights. Evaluation of racial and environmental justice impacts is one federally mandated example, where block level data on race/ethnic composition are used to test for differential impacts of social and environmental goods and bads, by various agencies including Environmental Protection Agency evaluations of environmental justice and US Forest Service environmental impact assessments. The Executive Order on *Advancing Racial Equity and Support for Underserved Communities Through the Federal Government* (01/20/2021), calls extra attention to the importance of accurate data on small area race/ethnic composition for meeting these demands.

- 4) CSAC recommends that the Census Bureau evaluate the impact of DP for racial equity uses, including Fair Housing Act and environmental justice, following the model of the redistricting evaluation (Wright and Iramata 2020).**

Census Bureau Response: The Census Bureau will look into the feasibility of adapting the Wright and Iramata analysis to the other use cases suggested in this recommendation.

VII) DP and Quality Metrics

CSAC finds that the quality metrics released for 2020 Census state level data have been extremely helpful in understanding the quality of that data. For example, these metrics reveal that in Louisiana 0.91 percent of all addresses were resolved through true count imputation (a rate 4 times the national average of 0.23 percent). Releasing this metric at the census tract level will help local planners understand the reliability of local area data, and where they may want to augment 2020 Census data with local administrative data for emergency response, road planning, and more. CSAC commends the Census Bureau for releasing a large number of quality metrics at the state level. The Bureau has already released census tract level self-response rates. CSAC cannot currently envision a justification for applying DP to census tract level metrics such as:

- Number of housing units with counts imputed
- Number of housing units enumerated by administrative data
- Number of housing units enumerated by proxy
- Number of NRFU housing units that were enumerated as nonexistent

- 1) CSAC recommends that the Census Bureau not apply DP to substate quality metrics that are being released to the public** -- either via the ASA task force or directly from the Census Bureau. Applying DP to the quality metrics could make them largely irrelevant, and will take part of the PLB away from important future data products. If the Bureau concludes that the quality metrics cannot be released without applying DP,

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

CSAC requests that the Bureau justify their decision by explaining how these metrics could be used in a reconstruction scenario.

Census Bureau Response: The Data Stewardship Executive Policy (DSEP) Committee will acknowledge this recommendation when considering requests to release substate quality metrics. However, if the frame for these statistics is the final 2020 Census frame and if the statistics interact with variables in official tabulations, it will be very difficult to exempt them from the differential privacy implementation in the 2020 DAS.

VIII) Evaluating Tradeoffs of Block/Block Group Data for Demographics and Housing Characteristics (DHC) and Other Upcoming Products

CSAC recognizes that the risk of disclosure is greatest for block level data cross-tabulated by detailed characteristics, such as age and race/ethnicity, that will be part of the DHC file.

- 1) CSAC recommends that the Census Bureau examine the utility of and explore tradeoffs of not releasing all the detailed tables in the DHC file at the block and potentially block group levels, if this would allow for more accuracy for small demographic groups (or housing vacancy, etc. variables) at other sub-state geographic levels.**

Census Bureau Response: The Census Bureau will take this recommendation under advisement as we turn our attention toward tuning the TopDown Algorithm for production of the 2020 Census Demographic and Housing Characteristics File scheduled for a later release.

It may be possible to use swapping, and potentially in combination with suppression for some small cells, in the DHC file for units at tract level and above. The goal would be to ensure that publicly released data are robust with a very clear indication of fitness for use. As the Bureau considers this strategy, they should engage with the user community to gain stronger fitness of use for other use cases.

IX) Federal Statistical Research Data Centers

Federal Statistical Research Data Centers (FSRDCs) are valuable resources built up over the last two decades. The idea that it is possible to access data in a controlled, restricted environment is important. Indeed, FSRDCs were an early answer to disclosure avoidance. They should be maintained.

- 1) CSAC recommends that the Bureau work on a plan to maintain FSRDC utility and access to high quality data.**

This would be especially important if the Bureau decides to restrict release of detailed block level data. Retaining this enclave approach, where credentialed users (such as researchers, government agencies or other decision-makers) can access data where

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

accuracy is not compromised by DP, would allow for the continuation of critical use cases such as the ability of policymakers to understand the impacts of past, present and future policies meant to improve the lives of Americans. Of course, research and conclusions stemming from FSRDC-based analyses would need to be assessed for privacy loss, with the likely application of DP to empirical findings before their approval for release.

Census Bureau Response: The Census Bureau accepts this recommendation. While the exact mechanisms and governance for FSRDC access to confidential 2020 Census data have not yet been determined, DSEP is committed to reserving 2020 Census privacy-loss budget to support FSRDC research.

- 2) CSAC recommends that the Bureau consider setting aside a portion of the PLB for such uses.**

Census Bureau Response: The Census Bureau accepts this recommendation. While the exact mechanisms and governance for FSRDC access to confidential 2020 Census data have not yet been determined, DSEP is committed to reserving 2020 Census privacy-loss budget to support FSRDC research.

X) Timeline for 2020 Census Product Releases

The Bureau's implementation of DP has followed an ambitious timeline. But many implications of DP implementation decisions are not yet fully understood. On the one hand, the risk for reconstruction attacks based on different levels of the PLB has not been fully quantified. On the other hand, the fitness for use of legal and regulatory uses of the data have not been examined in full, and risks of failing to produce sufficiently accurate data for these uses remain likely but not fully known. Certainly, Census data are used for a large range of funding, legal, and regulatory decisions at all levels of government. Stakeholders at the state level may be particularly aggrieved if decennial census products are insufficiently accurate for state decision-makers' legal, regulatory, and funding uses.

CSAC commends the Bureau for prioritizing research over speed of release in preparing the redistricting data. This is an important precedent for reducing risks (both in privacy loss and also in lack of fitness for use) before releasing data products.

- 1) CSAC recommends that the Census Bureau de-prioritize speed and prioritize performing the research necessary to understand and reduce the risk associated with privacy loss and insufficient fitness for use.**

Census Bureau Response: The Census Bureau accepts this recommendation. We are still finalizing the operational schedule for production of the 2020 Census Demographic and Housing Characteristics File, and in finalizing that schedule we are committed to including sufficient time for stakeholder engagement and evaluation to ensure both fitness-for-use and effective privacy protection for the data.

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Differential Privacy Meeting

While CSAC has provided information about some key use cases, many critical use cases are still unknown, and CSAC anticipates risks associated with releasing data products that are not sufficiently accurate for these use cases. A more complete use case catalog, as CSAC previously recommended, is still needed to identify and mitigate such risks.

- 2) CSAC recommends that the Bureau delay additional releases after the September redistricting file to allow sufficient time for developing the required new algorithms, testing the implications of alternative allocations of the PLB, assessing the risk of privacy loss from various epsilons, assessing risks of releasing data that is not fit-for-use (particularly for legal applications of decennial census data products), and developing demonstration products to inform users of the likely accuracy of the data.**

While CSAC recognizes that data users will be inconvenienced by further delays in releases of decennial data products, such delays will likely increase the accuracy of the resulting products while improving privacy protection, as the Bureau's techniques for developing and deploying DP are rapidly evolving. In addition, taking the time needed to do this work well will yield downstream benefits for other federal statistical agencies.

***Census Bureau Response:** The Census Bureau accepts this recommendation, though it may not necessarily require a delay. We are still finalizing the operational schedule for production of the 2020 Census Demographic and Housing Characteristics File, and in finalizing that schedule we are committed to including sufficient time for stakeholder engagement and evaluation to ensure both fitness-for-use and effective privacy protection for the data.*